



Application No. 10/806,092
Amendment dated September 12, 2006
Reply to Office Action of June 12, 2006

3

Docket No.: 1163-0500PUS1

AMENDMENTS TO THE CLAIMS

1. (Currently Amended) A document information processing apparatus comprising:
 - a plain document input unit for inputting a plain document;
 - a dictionary storage unit for storing a dictionary used for form element analysis and syntactic analysis;
 - a form element analyzer for performing a form element analysis on the plain document inputted from said plain document input unit by using the dictionary stored in said dictionary storage unit so as to decompose the plain document into tokens;
 - a syntax analyzer for analyzing a part of speech of each of the tokens obtained by said form element analyzer based on a syntax of said plain document so as to generate a structured document containing meaningful words;
 - a data storage unit for storing data used for a markup process;
 - an element refinement processing unit for performing the markup process of reading each of the meaningful words in the structured document and automatically adding data associated content to the structured document in association with each at least one of the meaningful words in order to generate a markup document, included in the structured document generated by said syntax analyzer and stored in said added content including data read from said data storage unit to each of the or data generated according to a determined attribute of the at least one of the meaningful words; words so as to generate a markup document; and
 - a markup document output unit for outputting the markup document generated by said element refinement processing unit.

2. (Currently Amended) The document information processing apparatus according to claim 1, wherein said apparatus further includes a text document input unit for generating a plain text document and sending ~~it~~ the generated plain text document to said plain document input unit.

3. (Currently Amended) The document information processing apparatus according to claim 1, wherein said apparatus further includes an external communication input unit for inputting a plain document from outside said apparatus by-via a communication device, and sends it~~the inputted plain document~~ to said plain document input unit.

4. (Currently Amended) The document information processing apparatus according to claim 1, wherein said apparatus further includes an external communication input unit for inputting a document on which a markup process is-has been performed from outside said apparatus by-via a communication device, and a markup document analyzer for analyzing and removing markups from the document on which the markup process is performed from inputted by said external communication input unit and removing markups from the document so as to generate a plain document and send it to be sent to said markup document input unit.

5. (Currently Amended) The document information processing apparatus according to Claim 1, wherein said apparatus further includes an external communication input unit for inputting an e-mail from outside said apparatus by-via a communication device, and an e-mail structure analyzer for analyzing a structure of the e-mail from said external communication input unit so as to generate a plain document, and for sending the generated plain document to a markup document input unit.

6. (Currently Amended) The document information processing apparatus according to Claim 2, wherein said apparatus further includes an external communication output unit for sending the markup document from said markup document output unit to an external device.
~~outside said apparatus.~~

7. (Currently Amended) The document information processing apparatus according to Claim 3, wherein said apparatus further includes an external communication output unit for sending the markup document from said markup document output unit to an external device.
~~outside said apparatus.~~

8. (Currently Amended) The document information processing apparatus according to Claim 4, wherein said apparatus further includes an external communication output unit for sending the markup document from said markup document output unit to an external device.
~~outside said apparatus.~~

9. (Currently Amended) The document information processing apparatus according to Claim 5, wherein said apparatus further includes an external communication output unit for sending the markup document from said markup document output unit to an external device.
~~outside said apparatus.~~

10. (Currently Amended) The document information processing apparatus according to Claim 1, wherein said element refinement processing unit includes an element relation analyzer for ~~associating a meaningful element with each of~~ automatically classifying at least one of the meaningful words ~~included in the structured document generated by said syntax analyzer as a~~ meaningful element, the meaningful element indicating a meaning or attribute of each of the corresponding meaningful word. words.

11. (Currently Amended) The document information processing apparatus according to Claim 10, wherein said apparatus further includes a proper name element processing unit for adding a markup to a proper name element indicating a proper name, and said element relation analyzer ~~starts-initiates operation of~~ said proper name element processing unit when the meaningful element is determined to be a proper name element.

12. (Currently Amended) The document information processing apparatus according to Claim 10, wherein said apparatus further includes a time element processing unit for adding a markup to a time element indicating a time, and said element relation analyzer ~~starts-initiates operation of~~ said time element processing unit when the meaningful element is determined to be a time element.

13. (Original) The document information processing apparatus according to Claim 12, wherein said apparatus further includes a time processing unit for computing a date corresponding to a time element, and said time element processing unit adds the date from said time processing unit, as a markup, to the time element.

14. (Currently Amended) The document information processing apparatus according to Claim 10, wherein said apparatus further includes a position element processing unit for adding a markup to a position element indicating a position, and said element relation analyzer starts initiates operation of said position element processing unit when the meaningful element is determined to be a position element.

15. (Original) The document information processing apparatus according to Claim 14, wherein said apparatus further includes a current position estimation unit for estimating a current position corresponding to the position element, and said position element processing unit adds the current position from said current position estimation unit, as a markup, to the position element.

16. (Currently Amended) The document information processing apparatus according to Claim 10, wherein said apparatus further includes an act element processing unit for adding a markup to an act element indicating an act, and said element relation analyzer starts-initiates operation of said act element processing unit when the meaningful element is determined to be an act element.

17. (New) The document information processing apparatus according to Claim 1, wherein

the element refinement processing unit determines an attribute of the at least one meaningful word, and adds a markup tag to the structured document identifying the attribute, the added content being associated with the markup tag.

18. (New) The document information processing apparatus according to Claim 17, wherein the added content is at least one of a definition, a Uniform Resource Identifier (URI) of a stored definition, or explanatory data relating to the attribute.

19. (New) The document information processing apparatus according to Claim 17, wherein the added content indicates a current position or date associated with the attribute.

20. (New) The document information processing apparatus according to Claim 17, wherein the determined attribute is classified as one of the following elements associated with the at least one meaningful word: a proper name element, a position element, a time element, and a verb element.

21. (New) The document information processing apparatus according to Claim 18, wherein the added content is obtained from a database entry in the data storage unit associated with the at least one meaningful word.